Tropical Plants

ARTICLE

Open Access

https://doi.org/10.48130/tp-0024-0030

Tropical Plants **2024**, 3: e027

Molecular insights into PIP aquaporins in tigernut (*Cyperus esculentus* L.), a Cyperaceae tuber plant

Authors

Zhi Zou[#*], Yujiao Zheng[#], Yanhua Xiao, Hongyan Liu, Jiaquan Huang[*], Yongguo Zhao[*]

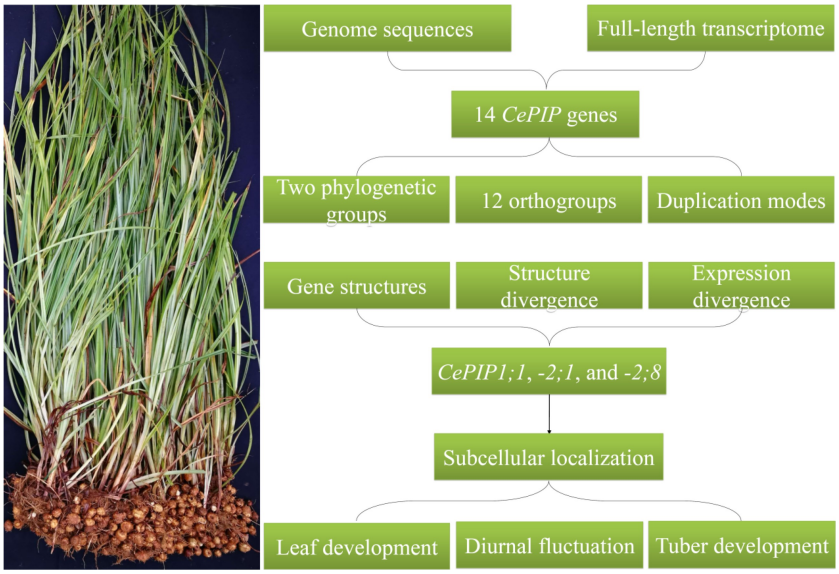
Correspondence

zouzhi@itbb.org.cn; jqhuang@hainanu.edu.cn; zhaoyongguo@gdupt.edu.cn

In Brief

This study presents the first comprehensive analysis of *PIP* genes in tigernut, which facilitates further elucidating the water balance mechanism in this oil-rich tuber plant.

Graphical abstract



Highlights

- Expansion of *CePIP* genes was mainly contributed by WGD and transposed/tandem duplications.
- *CePIP1;1*, *-2;1*, and *-2;8* have evolved to be the three dominant members.
- Their transcription in leaves is subjected to development and diurnal regulation.
- Their mRNA and protein abundances in tubers are positively correlated with the moisture content during tuber development.

Citation: Zou Z, Zheng Y, Xiao Y, Liu H, Huang J, et al. 2024. Molecular insights into PIP aquaporins in tigernut (*Cyperus esculentus* L.), a Cyperaceae tuber plant . *Tropical Plants* 3: e027 https://doi.org/10.48130/tp-0024-0030

Tropical Plants

ARTICLE

Open Access

https://doi.org/10.48130/tp-0024-0030

Tropical Plants **2024**, 3: e027

Molecular insights into PIP aquaporins in tigernut (*Cyperus esculentus* L.), a Cyperaceae tuber plant

Zhi Zou[1#*], Yujiao Zheng[1#], Yanhua Xiao[1], Hongyan Liu[2], Jiaquan Huang[2*] and Yongguo Zhao[3*]

[1] *National Key Laboratory for Tropical Crop Breeding, Institute of Tropical Biosciences and Biotechnology/Sanya Research Institute of Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, Hainan, P. R. China*

[2] *School of Tropical Agriculture and Forestry, Hainan University, Haikou 571101, Hainan, P. R. China*

[3] *College of Biology and Food Engineering, Guangdong University of Petrochemical Technology, Maoming 525000, Guangdong, P. R. China*

[#] Authors contributed equally: Zhi Zou, Yujiao Zheng

* Corresponding authors, E-mail: zouzhi@itbb.org.cn; jqhuang@hainanu.edu.cn; zhaoyongguo@gdupt.edu.cn

Abstract

Plasma membrane intrinsic proteins (PIPs) constitute a subfamily of aquaporin that mediate the efficient water transport across the cell membrane. In the present study, the first genome-wide characterization of *PIP* genes in tigernut (*Cyperus esculentus* L.), a Cyperaceae plant producing oil in underground tubers, is presented. A high number of 14 *PIP* genes representing two phylogenetic groups or 12 orthogroups were identified from the tigernut genome. Comparative genomics analyses revealed that the expansion of *CePIP* genes was mainly contributed by whole-genome duplication and transposed/tandem duplications, and some of them are lineage or even species-specific. Expression analyses indicated that *CePIP1;1*, *-2;1*, and *-2;8* are three dominant members, which are constitutively expressed in most tissues including leaf and tuber. Whereas transcription of *CePIP1;1*, *-2;1*, and *-2;8* in leaves is subjected to development and diurnal regulation, their mRNA/protein abundances in tubers are positively correlated with the moisture content during tuber development. Moreover, when transiently overexpressed in *Nicotiana benthamiana* leaves, all three proteins were shown to be located in the cell membrane. These findings provide valuable information for further functional analysis and genetic improvement in tigernut.

Citation: Zou Z, Zheng Y, Xiao Y, Liu H, Huang J, et al. 2024. Molecular insights into PIP aquaporins in tigernut (*Cyperus esculentus* L.), a Cyperaceae tuber plant . *Tropical Plants* 3: e027 https://doi.org/10.48130/tp-0024-0030

Introduction

Aquaporins (AQPs) constitute a large family of transmembrane channel proteins that function as regulators of intracellular and intercellular water flow[1,2]. Since their first discovery in the 1990s, AQPs have been found not only in three domains of life, i.e., bacteria, eukaryotes, and archaea, but also in viruses[3,4]. Each AQP monomer is composed of an internal repeat of three transmembrane helices (i.e., TM1–TM6) as well as two half helixes that are formed by loop B (LB) and LE through dipping into the membrane[5]. The dual Asn-Pro-Ala (NPA) motifs that are located at the N-terminus of two half helixes act as a size barrier of the pore *via* creating an electrostatic repulsion of protons, whereas the so-called aromatic/arginine (ar/R) selectivity filter (i.e., H2, H5, LE1, and LE2) determines the substrate specificity by rendering the pore constriction site diverse in both size and hydrophobicity[5–9]. Based on sequence similarity, AQPs in higher plants could be divided into five subfamilies, i.e., plasma membrane intrinsic protein (PIP), tonoplast intrinsic protein (TIP), NOD26-like intrinsic protein (NIP), X intrinsic protein (XIP), and small basic intrinsic protein (SIP)[10–17]. Among them, PIPs, which are typically localized in the cell membrane, are most conserved and play a central role in controlling plant water status[12,18–22]. Among two phylogenetic groups present in the PIP subfamily, PIP1 possesses a relatively longer N-terminus and PIP2 features an extended C-terminus with one or more conserved S residues for phosphorylation modification[5,15,17].

Tigernut (*Cyperus esculentus* L.), which belongs to the Cyperaceae family within Poales, is a novel and promising herbaceous C_4 oil crop with wide adaptability, large biomass, and short life period[23–27]. Tigernut is a unique species accumulating up to 35% oil in the underground tubers[28–30], which are developed from stolons and the process includes three main stages, i.e., initiation, swelling, and maturation[31–33]. Water is essential for tuber development and tuber moisture content maintains a relatively high level of approximately 85% until maturation when a significant drop to about 45% is observed[28,32]. Thereby, uncovering the mechanism of tuber water balance is of particular interest. Despite crucial roles of PIPs in the cell water balance, to date, their characterization in tigernut is still in the infancy[21]. The recently available genome and transcriptome datasets[31,33,34] provide an opportunity to address this issue.

In this study, a global characterization of *PIP* genes was conducted in tigernut, including gene localizations, gene structures, sequence characteristics, and evolutionary patterns. Moreover, the correlation of *CePIP* mRNA/protein abundance with water content during tuber development as well as subcellular localizations were also investigated, which facilitated further elucidating the water balance mechanism in this special species.

Materials and methods

Datasets and identification of *PIP* genes

PIP genes reported in Arabidopsis (*Arabidopsis thaliana*)[10] and rice (*Oryza sativa*)[11] were respectively obtained from TAIR11 (www.arabidopsis.org) and RGAP7 (http://rice.uga.edu), and detailed information is shown in Supplemental Table S1. Their protein sequences were used as queries for tBLASTn[35] (*E*-value, 1e–10) search of the full-length tigernut transcriptome and genome sequences that were accessed from CNGBdb (https://db.cngb.org/search/assembly/CNA0051961)[31,34]. RNA sequencing (RNA-seq) reads that are available in NCBI (www.ncbi.nlm.nih.gov/sra) were also adopted for gene structure revision as described before[13], and presence of the conserved MIP (major intrinsic protein, Pfam accession number PF00230) domain in candidates was confirmed using MOTIF Search (www.genome.jp/tools/motif). To uncover the origin and evolution of *CePIP* genes, a similar approach was also employed to identify homologs from representative plant species, i.e., *Carex cristatella* (v1, Cyperaceae)[36], *Rhynchospora breviuscula* (v1, Cyperaceae)[37], and *Juncus effusus* (v1, Juncaceae)[37], whose genome sequences were accessed from NCBI (www.ncbi.nlm.nih.gov). Gene structure of candidates were displayed using GSDS 2.0 (http://gsds.gao-lab.org), whereas physiochemical parameters of deduced proteins were calculated using ProtParam (http://web.expasy.org/protparam). Subcellular localization prediction was conducted using WoLF PSORT (www.genscript.com/wolf-psort.html).

Sequence alignment, phylogenetic analysis, and identification of conserved motifs

Nucleotide and protein multiple sequence alignments were respectively conducted using ClustalW and MUSCLE implemented in MEGA6[38] with default parameters, and phylogenetic tree construction was carried out using MEGA6 with the maximum likelihood method and bootstrap of 1,000 replicates. Systematic names of *PIP* genes were assigned with two italic letters denoting the source organism and a progressive number based on sequence similarity. Conserved motifs were identified using MEME Suite 5.5.3 (https://meme-suite.org/tools/meme) with optimized parameters as follows: Any number of repetitions, maximum number of 15 motifs, and a width of 6 and 250 residues for each motif. TMs and conserved residues were identified using homology modeling and sequence alignment with the structure resolved spinach (*Spinacia oleracea*) SoPIP2;1[5].

Synteny analysis, definition of orthogroups, and calculation of the evolutionary rate

Synteny analysis was conducted using TBtools-II[39] as described previously[40], where the parameters were set as *E*-value of 1e-10 and BLAST hits of 5. Duplication modes were identified using the DupGen_finder pipeline[41], and Ks (synonymous substitution rate) and Ka (nonsynonymous substitution rate) of duplicate pairs were calculated using codeml in the PAML package[42]. Orthologs between different species were identified using InParanoid[43] and information from synteny analysis, and orthogroups (OGs) were assigned only when they were present in at least two species examined.

Plant materials

Plant materials used for gene cloning, qRT-PCR analysis, and 4D-parallel reaction monitoring (4D-PRM)-based protein quantification were derived from a tigernut variety Reyan3[31], and plants were grown in a greenhouse as described previously[25]. For expression profiling during leaf development, three representative stages, i.e., young, mature, and senescing, were selected and the chlorophyll content was checked using SPAD-502Plus (Konica Minolta, Shanghai, China) as previously described[44]. Young and senescing leaves are yellow in appearance, and their chlorophyll contents are just half of that of mature leaves that are dark green. For diurnal fluctuation regulation, mature leaves were sampled every 4 h from the onset of light at 8 a.m. For gene regulation during tuber development, fresh tubers at 1, 5, 10, 15, 20, 25, and 35 d after tuber initiation (DAI) were collected as described previously[32]. All samples with three biological replicates were quickly frozen with liquid nitrogen and stored at −80 °C for further use. For subcellular localization analysis, tobacco (*Nicotiana benthamiana*) plants were grown as previously described[20].

Gene expression analysis

Tissue-specific expression profiles of *CePIP* genes were investigated using Illumina RNA-seq samples (150 bp paired-end reads) with three biological replicates for young leaf, mature leaf, sheath of mature leaf, shoot apex, root, rhizome, and three stages of developmental tuber (40, 85, and 120 d after sowing (DAS)), which are under the NCBI accession number of PRJNA703731. Raw sequence reads in the FASTQ format were obtained using fastq-dump, and quality control was performed using fastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc). Read mapping was performed using HISAT2 (v2.2.1, https://daehwankimlab.github.io/hisat2), and relative gene expression level was presented as FPKM (fragments per kilobase of exon per million fragments mapped)[45].

For qRT-PCR analysis, total RNA extraction and synthesis of the first-strand cDNA were conducted as previously described[24]. Primers used in this study are shown in Supplemental Table S2, where *CeUCE2* and *CeTIP41*[25,33] were employed as two reference genes. PCR reaction in triplicate for each biological sample was carried out using the SYBR-green Mix (Takara) on a Real-time Thermal Cycler Type 5100 (Thermal Fisher Scientific Oy). Relative gene abundance was estimated with the $2^{-\Delta\Delta C_T}$ method and statistical analysis was performed using SPSS Statistics 20 as described previously[13].

Protein quantification

Raw proteomic data for tigernut roots, leaves, freshly harvested, dried, rehydrated for 48 h, and sprouted tubers were downloaded from ProteomeXchange/PRIDE (www.proteomexchange.org, PXD021894, PXD031123, and PXD035931), which were further analyzed using Maxquant (v1.6.15.0, www.maxquant.org). Three dominant members, i.e., CePIP1;1, -2;1, and -2;8, were selected for 4D-PRM quantification analysis, and related unique peptides are shown in Supplemental Table S3. Protein extraction, trypsin digestion, and LC-MS/MS analysis were conducted as described previously[46].

Subcellular localization analysis

For subcellular localization analysis, the coding region (CDS) of *CePIP1;1*, *-2;1*, and *-2;8* were cloned into *p*NC-Cam1304-SubN *via* Nimble Cloning as described before[30]. Then, recombinant plasmids were introduced into *Agrobacterium tumefaciens* GV3101 with the helper plasmid *p*Soup-*P19* and infiltration of 4-week-old tobacco leaves were performed as previously described[20]. For subcellular localization analysis, the plasma

membrane marker HbPIP2;3-RFP[22] was co-transformed as a positive control. Fluorescence observation was conducted using confocal laser scanning microscopy imaging (Zeiss LMS880, Germany): The wavelength of laser-1 was set as 730 nm for RFP observation, where the fluorescence was excited at 561 nm; the wavelength of laser-2 was set as 750 nm for EGFP observation, where the fluorescence was excited at 488 nm; and the wavelength of laser-3 was set as 470 nm for chlorophyll autofluorescence observation, where the fluorescence was excited at 633 nm.

Results

Characterization of 14 *PIP* genes in tigernut

As shown in Table 1, a total of 14 *PIP* genes were identified from eight tigernut scaffolds (Scfs). The CDS length varies from 831 to 882 bp, putatively encoding 276–293 amino acids (AA) with a molecular weight (MW) of 29.16–31.59 kilodalton (kDa). The theoretical isoelectric point (pI) varies from 7.04 to 9.46, implying that they are all alkaline. The grand average of hydropathicity (GRAVY) is between 0.344 and 0.577, and the aliphatic index (II) ranges from 94.57 to 106.90, which are consistent with the hydrophobic characteristic of AQPs[47]. As expected, like SoPIP2;1, all CePIPs include six TMs, two typical NPA motifs, the invariable ar/R filter F-H-T-R, five conserved Froger's positions Q/M-S-A-F-W, and two highly conserved residues corresponding to H^{193} and L^{197} in SoPIP2;1 that were proven to be involved in gating[5,48], though the H→F variation was found in CePIP2;9, -2;10, and -2;11 (Supplemental Fig. S1). Moreover, two S residues, corresponding to S^{115} and S^{274} in SoPIP2;1[5], respectively, were also found in the majority of CePIPs (Supplemental Fig. S1), implying their posttranslational regulation by phosphorylation.

To uncover the evolutionary relationships, an unrooted phylogenetic tree was constructed using the full-length protein sequences of CePIPs together with 11 OsPIPs and 13 AtPIPs. As shown in Fig. 1a, these proteins were clustered into two main groups, corresponding to PIP1 and PIP2 as previously defined[10,49], and each appears to have evolved into several subgroups. Compared with PIP1s, PIP2s possess a relatively shorter N-terminal but an extended C-terminal with one conserved S residue (Supplemental Fig. S1). Interestingly, a high number of gene repeats were detected, most of which seem to be species-specific, i.e., *AtPIP1;1/-1;2/-1;3/-1;4/-1;5*, *AtPIP2;1/-2;2/-2;3/-2;4/-2;5/-2;6*, *AtPIP2;7/-2;8*, *OsPIP1;1/-1;2/-1;3*, *OsPIP2;1/-2;4/-2;5*, *OsPIP2;2/-2;3*, *CePIP1;1/-1;2*, *CePIP2;2/-2;3*, *CePIP2;4/-2;5/-2;6/-2;7*, and *CePIP2;9/-2;10/-2;11*, reflecting the occurrence of more than one lineage-specific whole-genome duplications (WGDs) after their divergence[50,51]. In Arabidopsis that experienced three WGDs (i.e. γ, β, and α) after the split with the monocot clade[52], *AtPIP1;5* in the PIP1 group first gave rise to *AtPIP1;1 via* the γ WGD shared by all core eudicots[50], which latter resulted in *AtPIP1;3*, *-1;4*, and *-1;2 via* β and α WGDs; *AtPIP2;1* in the PIP2 group first gave rise to *AtPIP2;6 via* the γ WGD, and they latter generated *AtPIP2;2*, and *-2;5 via* the α WGD (Supplemental Table S1). In rice, which also experienced three WGDs (i.e. τ, σ, and ρ) after the split with the eudicot clade[51], *OsPIP1;2* and *-2;3* generated *OsPIP1;1* and *-2;2 via* the Poaceae-specific ρ WGD, respectively. Additionally, tandem, proximal, transposed and dispersed duplications also played a role on the gene expansion in these two species (Supplemental Table S1).

Analysis of gene structures revealed that all *CePIP* and *AtPIP* genes possess three introns and four exons in the CDS, in contrast to the frequent loss of certain introns in rice, including *OsPIP1;2*, *-1;3*, *-2;1*, *-2;3*, *-2;4*, *-2;5*, *-2;6*, *-2;7*, and *-2;8* (Fig. 1b). The positions of three introns are highly conserved, which are located in sequences encoding LB (three residues before the first NPA), LD (one residue before the conserved L involved in gating), and LE (18 residues after the second NPA), respectively (Supplemental Fig. S1). The intron length of *CePIP* genes is highly variable, i.e., 109–993 bp, 115–1745 bp, and 95–866 bp for three introns, respectively. By contrast, the exon length is relatively less variable: Exons 2 and 3 are invariable with 296 bp and 141 bp, respectively, whereas Exons 1 and 4 are of 277–343 bp and 93–132 bp, determining the length of N- and C-terminus of PIP1 and PIP2, respectively (Fig. 1b). Correspondingly, their protein structures were shown to be highly conserved, and six (i.e., Motifs 1–6) out of 15 motifs identified are broadly present. Among them, Motif 3, -2, -6, -1, and -4 constitute the conserved MIP domain. In contrast to a single Motif 5 present in most PIP2s, all PIP1s possess two sequential copies of Motif 5, where the first one is located at the extended N-terminal. In CePIP2;3 and OsPIP2;7, Motif 5 is replaced by Motif 13; in

Table 1. Fourteen *PIP* genes identified in *C. esculentus*.

Gene name	Locus	Position	Intron no.	AA	MW (kDa)	pI	GRAVY	AI	TM	MIP
CePIP1;1	CESC_15147	Scf9:2757378..2759502(–)	3	288	30.76	8.82	0.384	95.28	6	47..276
CePIP1;2	CESC_04128	Scf4:3806361..3807726(–)	3	291	31.11	8.81	0.344	95.95	6	46..274
CePIP1;3	CESC_15950	Scf54:5022493..5023820(+)	3	289	31.06	8.80	0.363	94.57	6	49..278
CePIP2;1	CESC_15350	Scf9:879960..884243(+)	3	288	30.34	8.60	0.529	103.02	6	33..269
CePIP2;2	CESC_00011	Scf30:4234620..4236549(+)	3	293	31.59	9.27	0.394	101.57	6	35..268
CePIP2;3	CESC_00010	Scf30:4239406..4241658(+)	3	291	30.88	9.44	0.432	98.97	6	31..266
CePIP2;4	CESC_05080	Scf46:307799..309544(+)	3	285	30.44	7.04	0.453	100.32	6	28..265
CePIP2;5	CESC_05079	Scf46:312254..314388(+)	3	286	30.49	7.04	0.512	101.68	6	31..268
CePIP2;6	CESC_05078	Scf46:316024..317780(+)	3	288	30.65	7.68	0.475	103.06	6	31..268
CePIP2;7	CESC_05077	Scf46:320439..322184(+)	3	284	30.12	8.55	0.500	100.00	6	29..266
CePIP2;8	CESC_14470	Scf2:4446409..4448999(+)	3	284	30.37	8.30	0.490	106.90	6	33..263
CePIP2;9	CESC_02223	Scf1:2543928..2545778(–)	3	283	30.09	9.46	0.533	106.47	6	31..262
CePIP2;10	CESC_10007	Scf27:1686032..1688010(–)	3	276	29.16	9.23	0.560	106.05	6	26..256
CePIP2;11	CESC_10009	Scf27:1694196..1696175(–)	3	284	29.71	9.10	0.577	105.49	6	33..263

AA: amino acid; AI: aliphatic index; GRAVY: grand average of hydropathicity; kDa: kilodalton; MIP: major intrinsic protein; MW: molecular weight; pI: isoelectric point; PIP: plasma membrane intrinsic protein; Scf: scaffold; TM: transmembrane helix.

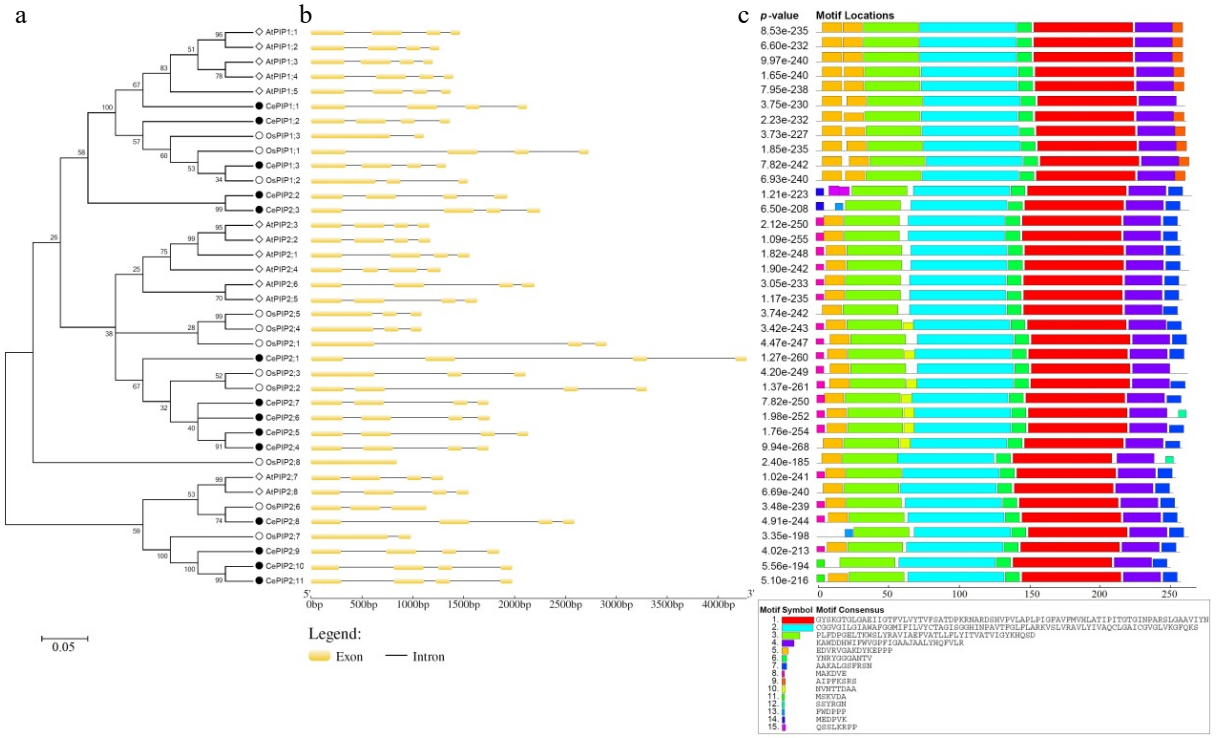


Fig. 1 Structural and phylogenetic analysis of PIPs in *C. esculentus*, *O. sativa*, and *A. thaliana*. (a) Shown is an unrooted phylogenetic tree resulting from full-length PIPs with MEGA6 (maximum likelihood method and bootstrap of 1,000 replicates), where the distance scale denotes the number of amino acid substitutions per site. (b) Shown are the exon-intron structures. (c) Shown is the distribution of conserved motifs among PIPs, where different motifs are represented by different color blocks as indicated and the same color block in different proteins indicates a certain motif. (At: *A. thaliana*; Ce: *C. esculentus*; PIP: plasma membrane intrinsic protein; Os: *O. sativa*).

CePIP2;2, it is replaced by two copies of Motif 15; and no significant motif was detected in this region of CePIP2;10. PIP1s and PIP2s usually feature Motif 9 and -7 at the C-terminal, respectively, though it is replaced by Motif 12 in CePIP2;6 and OsPIP2;8. PIP2s usually feature Motif 8 at the N-terminal, though it is replaced by Motif 14 in CePIP2;2 and -2;3 or replaced by Motif 11 in CePIP2;10 and -2;11 (Fig. 1c).

Gene localization and synteny analyses reveal complex evolution patterns of *CePIP* genes

As shown in Fig. 2a, gene localization of *CePIPs* revealed three gene clusters, i.e., *CePIP2;2/-2;3* on Scf30, *CePIP2;4/-2;5/-2;6/-2;7* on Scf46, and *CePIP2;10/-2;11* on Scf27, which were defined as tandem repeats for their high sequence similarities and neighboring locations. The nucleotide identities of these duplicate pairs vary from 70.5% to 91.2%, and the Ks values range from 0.0971 to 1.2778 (Table 2), implying different time of their birth. According to intra-species synteny analysis, two duplicate pairs, i.e., *CePIP1;1/-1;2* and *CePIP2;2/-2;4*, were shown to be located within syntenic blocks (Fig. 2b) and thus were defined as WGD repeats. Among them, *CePIP1;1/-1;2* possess a comparable Ks value to *CePIP2;2/-2;3*, *CePIP1;1/-1;3*, and *CePIP2;4/-2;8* (1.2522 vs 1.2287–1.2778), whereas *CePIP2;2/-2;4* harbor a relatively higher Ks value of 1.5474 (Table 2), implying early origin or fast evolution of the latter. While *CePIP1;1/-1;3* and *CePIP2;1/-2;8* were characterized as transposed repeats, *CePIP2;1/-2;2*, *CePIP2;9/-2;10*, and *CePIP2;8/-2;10* were characterized as dispersed repeats (Fig. 2a). The Ks values of three dispersed repeats vary from 0.8591 to 3.0117 (Table 2), implying distinct times of origin.

According to inter-species syntenic analysis, six out of 14 *CePIP* genes were shown to have syntelogs in rice, including 1:1, 1:2, and 2:2 (i.e. *CePIP1;1* vs *OsPIP1;3*, *CePIP1;3* vs *OsPIP1;2/-1;1*, *CePIP2;1* vs *OsPIP2;4*, *CePIP2;2/-2;4* vs *OsPIP2;3/-2;2*, and *CePIP2;8* vs *OsPIP2;6*), in striking contrast to a single one found in Arabidopsis (i.e. *CePIP1;2* vs *AtPIP1;2*). Correspondingly, only *OsPIP1;2* in rice was shown to have syntelogs in Arabidopsis, i.e., *AtPIP1;3* and *-1;4* (Fig. 2b). These results are consistent with their taxonomic relationships that tigernut and rice are closely related[50,51], and also imply lineage-specific evolution after their divergence.

Comparative genomics analyses reveals lineage-specific evolution of *PIP* genes in Cyperaceae

As described above, phylogenetic and syntenic analyses showed that the last common ancestor of tigernut and rice is more likely to possess only two PIP1s and three PIP2s. However, it is not clear whether the gene expansion observed in tigernut is species-specific or Cyperaceae-specific. To address this issue, recently available genomes were used to identify *PIP* subfamily genes from *C. cristatella*, *R. breviuscula*, and *J. effuses*, resulting in 15, 13, and nine members, respectively. Interestingly, in contrast to a high number of tandem repeats found in Cyperaceae species, only one pair of tandem repeats (i.e., *JePIP2;3* and *-2;4*) were identified in *J. effusus*, a close outgroup species to Cyperaceae in the Juncaceae family[36,37]. According to homologous analysis, a total of 12 orthogroups were identified, where *JePIP* genes belong to PIP1A (*JePIP1;1*), PIP1B (*JePIP1;2*), PIP1C (*JePIP1;3*), PIP2A (*JePIP2;1*), PIP2B (*JePIP2;2*), PIP2F (*JePIP2;3* and *-2;4*), PIP2G (*JePIP2;5*), and PIP2H (*JePIP2;6*) (Table

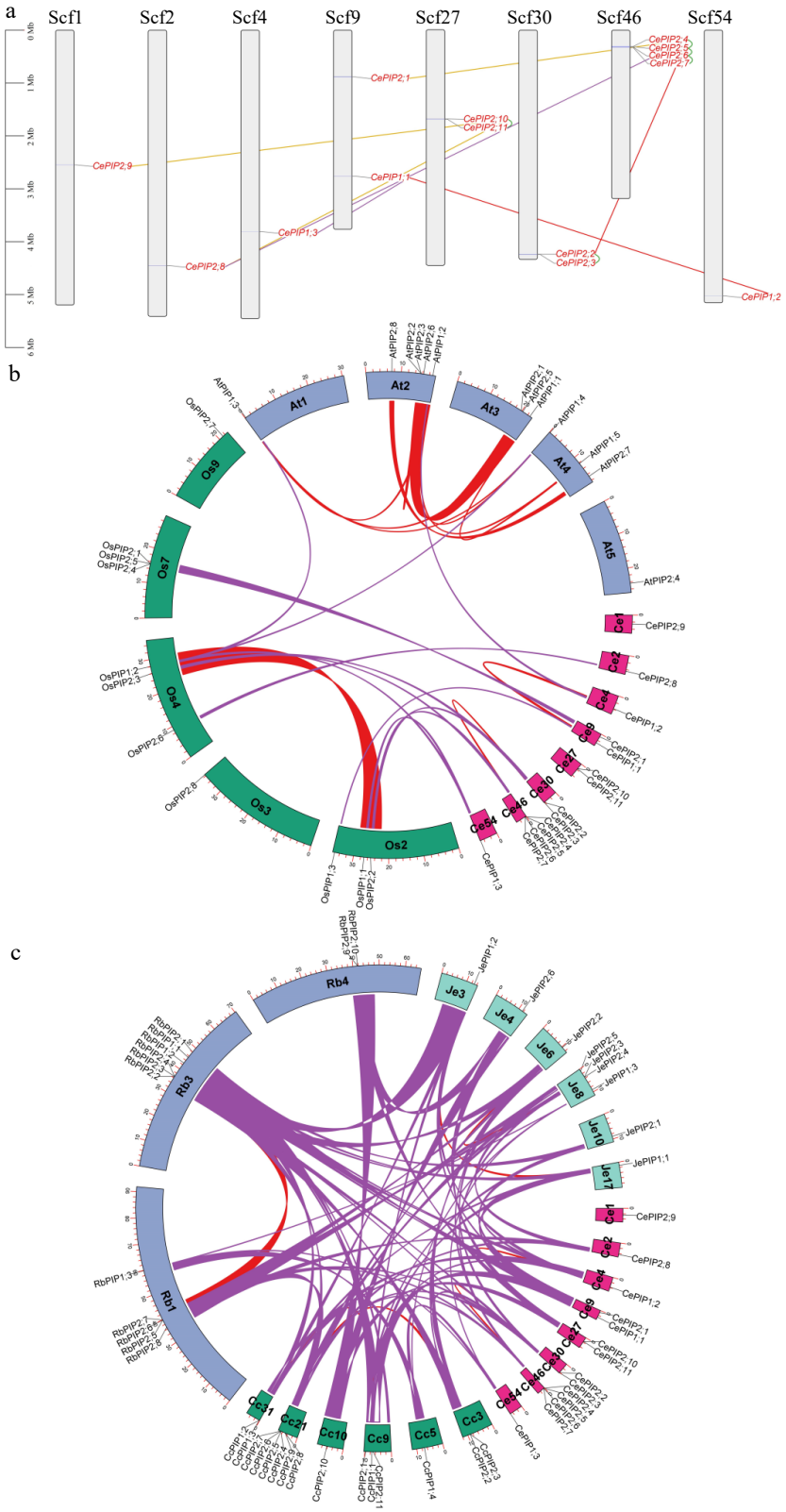


Fig. 2 Duplication events of *CePIP* genes and synteny analysis within and between *C. esculentus*, *O. sativa*, and *A. thaliana*. (a) Duplication events detected in tigernut. Serial numbers are indicated at the top of each scaffold, and the scale is in Mb. Duplicate pairs identified in this study are connected using lines in different colors, i.e., tandem (shown in green), transposed (shown in purple), dispersed (shown in gold), and WGD (shown in red). (b) Synteny analysis within and between *C. esculentus*, *O. sativa*, and *A. thaliana*. (c) Synteny analysis within and between *C. esculentus*, *C. cristatella*, *R. breviuscula*, and *J. effusus*. Shown are *PIP*-encoding chromosomes/scaffolds and only syntenic blocks that contain *PIP* genes are marked, i.e., red and purple for intra- and inter-species, respectively. (At: *A. thaliana*; Cc: *C. cristatella*; Ce: *C. esculentus*; Je: *J. effusus*; Mb: megabase; PIP: plasma membrane intrinsic protein; Os: *O. sativa*; Rb: *R. breviuscula*; Scf: scaffold; WGD: whole-genome duplication).

Table 2. Sequence identity and evolutionary rate of homologous *PIP* gene pairs identified in *C. esculentus*. Ks and Ka were calculated using PAML.

Duplicate 1	Duplicate 2	Identity (%)	Ka	Ks	Ka/Ks
CePIP1;1	*CePIP1;3*	78.70	0.0750	1.2287	0.0610
CePIP1;2	*CePIP1;1*	77.20	0.0894	1.2522	0.0714
CePIP2;1	*CePIP2;4*	74.90	0.0965	1.7009	0.0567
CePIP2;3	*CePIP2;2*	70.50	0.1819	1.2778	0.1424
CePIP2;4	*CePIP2;2*	66.50	0.2094	1.5474	0.1353
CePIP2;5	*CePIP2;4*	87.30	0.0225	0.4948	0.0455
CePIP2;6	*CePIP2;5*	84.90	0.0545	0.5820	0.0937
CePIP2;7	*CePIP2;6*	78.70	0.0894	1.0269	0.0871
CePIP2;8	*CePIP2;4*	72.90	0.1401	1.2641	0.1109
CePIP2;9	*CePIP2;10*	76.40	0.1290	0.8591	0.1502
CePIP2;10	*CePIP2;8*	64.90	0.2432	3.0117	0.0807
CePIP2;11	*CePIP2;10*	91.20	0.0562	0.0971	0.5783

Ce: *C. esculentus*; Ka: nonsynonymous substitution rate; Ks: synonymous substitution rate; PIP: plasma membrane intrinsic protein.

3). Further intra-species syntenic analysis revealed that *JePIP1;1/-1;2* and *JePIP2;2/-2;3* are located within syntenic blocks, which is consistent with *CePIP1;1/-1;2*, *CePIP2;2/-2;4*, *CcPIP1;1/-1;2*, *CcPIP2;3/-2;4*, *RbPIP1;1/-1;2*, and *RbPIP2;2/-2;5* (Fig. 2c), implying that PIP1A/PIP1B and PIP2B/PIP2D were derived from WGDs occurred sometime before Cyperaceae-Juncaceae divergence. After the split with Juncaceae, tandem duplications frequently occurred in Cyperaceae, where PIP2B/PIP2C and PIP2D/PIP2E/PIP2F retain in most Cyperaceae plants examined in this study. By contrast, species-specific expansion was also observed, i.e., *CePIP2;4/-2;5*, *CePIP2;10/-2;11*, *CcPIP1;2/-1;3*, *CcPIP2;4/-2;5*, *CcPIP2;8/-2;9*, *CcPIP2;10/-2;11*, *RbPIP2;3/-2;4*, and *RbPIP2;9/-2;10* (Table 3 & Fig. 2c).

Tissue-specific expression analysis of *CePIP* genes identifies three dominant members

Tissue-specific expression profiles of *CePIP* genes were investigated using transcriptome data available for young leaf, mature leaf, sheath, root, rhizome, shoot apex, and tuber. As shown in Fig. 3a, *CePIP* genes were mostly expressed in roots, followed by sheaths, moderately in tubers, young leaves, rhizomes, and mature leaves, and lowly in shoot apexes. In most tissues, *CePIP1;1*, *-2;1*, and *-2;8* represent three dominant members that contributed more than 90% of total transcripts. By contrast, in rhizome, these three members occupied about 80% of total transcripts, which together with *CePIP1;3* and *-2;4* contributed up to 96%; in root, *CePIP1;1*, *-1;3*, *-2;4*, and *-2;7* occupied about 84% of total transcripts, which together with *CePIP2;1* and *-2;8* contributed up to 94%. According to their expression patterns, *CePIP* genes could be divided into five main clusters: Cluster I includes *CePIP1;1*, *-2;1*, and *-2;8* that were constitutively and highly expressed in all tissues examined; Cluster II includes *CePIP2;2*, *-2;9*, and *-2;10* that were lowly expressed in all tested tissues; Cluster III includes *CePIP1;2* and *-2;11* that were preferentially expressed in young leaf and sheath; Cluster IV includes *CePIP1;3* and *-2;4* that were predominantly expressed in root and rhizome; and Cluster V includes remains that were typically expressed in root (Fig. 3a). Collectively, these results imply expression divergence of most duplicate pairs and three members (i.e. *CePIP1;1*, *-2;1*, and *-2;8*) have evolved to be constitutively co-expressed in most tissues.

Expression analysis reveals that *CePIP* transcripts in leaves are subjected to development regulation

As shown in Fig. 3a, compared with young leaves, transcriptome profiling showed that *CePIP1;2*, *-2;3*, *-2;7*, *-2;8*, and *-2;11* were significantly down-regulated in mature leaves, whereas *CePIP1;3* and *-2;1* were up-regulated. To confirm the results, three dominant members, i.e., *CePIP1;1*, *-2;1*, and *-2;8*, were selected for qRT-PCR analysis, which includes three representative stages, i.e., young, mature, and senescing leaves. As shown in Fig. 3b, in contrast to *CePIP2;1* that exhibited a bell-like expression pattern peaking in mature leaves, transcripts of both *CePIP1;1* and *-2;8* gradually decreased during leaf development. These results were largely consistent with transcriptome profiling, and the only difference is that *CePIP1;1* was significantly down-regulated in mature leaves relative to young leaves. However, this may be due to different experiment conditions used, i.e., greenhouse vs natural conditions.

qRT-PCR analysis reveals that the transcription of *CePIP1;1*, *-2;1*, and *-2;8* in mature leaves is subjected to diurnal regulation

Diurnal fluctuation expression patterns of *CePIP1;1*, *-2;1*, and *-2;8* were also investigated in mature leaves and results are

Table 3. Twelve proposed orthogroups based on comparison of representative plant species.

Orthogroup	*C. esculentus*	*C. cristatella*	*R. breviuscula*	*J. effusus*	*O. sativa*	*A. thaliana*
PIP1A	CePIP1;1	CcPIP1;1	RbPIP1;1	JePIP1;1	OsPIP1;3	AtPIP1;1, AtPIP1;2, AtPIP1;3, AtPIP1;4, AtPIP1;5
PIP1B	CePIP1;2	CcPIP1;2, CcPIP1;3	RbPIP1;2	JePIP1;2	–	–
PIP1C	CePIP1;3	CcPIP1;4	RbPIP1;3	JePIP1;3	OsPIP1;1, OsPIP1;2	–
PIP2A	CePIP2;1	CcPIP2;1	RbPIP2;1	JePIP2;1	OsPIP2;1, OsPIP2;4, OsPIP2;5	AtPIP2;1, AtPIP2;2, AtPIP2;3, AtPIP2;4, AtPIP2;5, AtPIP2;6
PIP2B	CePIP2;2	CcPIP2;2	RbPIP2;2	JePIP2;2	OsPIP2;2, OsPIP2;3	–
PIP2C	CePIP2;3	CcPIP2;3	RbPIP2;3, RbPIP2;4	–	–	–
PIP2D	CePIP2;4, CePIP2;5	CcPIP2;4, CcPIP2;5	RbPIP2;5	–	–	–
PIP2E	CePIP2;5	CcPIP2;5	RbPIP2;6	–	–	–
PIP2F	CePIP2;6	CcPIP2;6	–	–	–	–
PIP2G	CePIP2;7	CcPIP2;7	RbPIP2;7	JePIP2;3, JePIP2;4	–	–
PIP2H	CePIP2;8	CcPIP2;8, CcPIP2;9	RbPIP2;8	JePIP2;5	OsPIP2;6	AtPIP2;7, AtPIP2;8
PIP2I	CePIP2;9, CePIP2;10, CePIP2;11	CcPIP2;10, CcPIP2;11	RbPIP2;9, RbPIP2;10	JePIP2;6	OsPIP2;7, OsPIP2;8	–

At: *A. thaliana*; Cc: *C. cristatella*; Ce: *C. esculentus*; Je: *J. effuses*; Os: *O. sativa*; Rb: *R. breviuscula*; PIP: plasma membrane intrinsic protein.

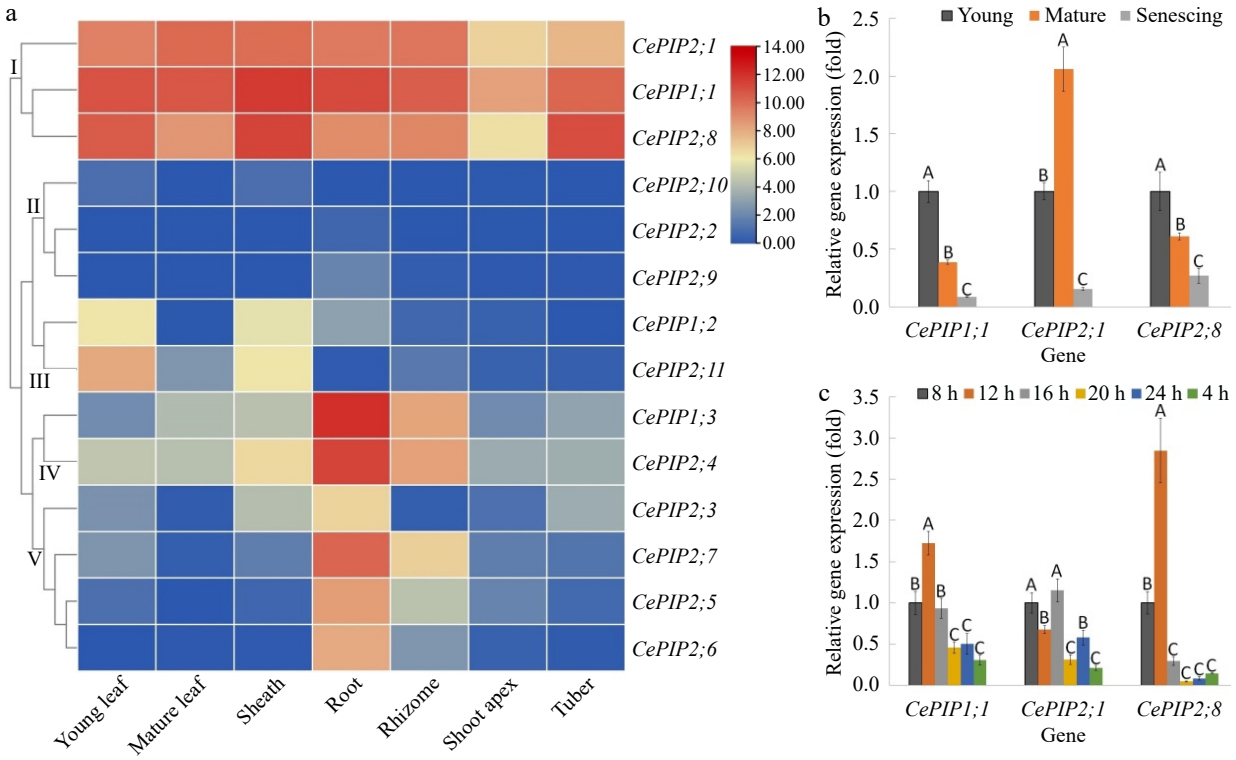


Fig. 3 Expression profiles of *CePIP* genes in various tissues, different stages of leaf development, and mature leaves of diurnal fluctuation. (a) Tissue-specific expression profiles of 14 *CePIP* genes. The heatmap was generated using the R package implemented with a row-based standardization. Color scale represents FPKM normalized $\log_2$ transformed counts, where blue indicates low expression and red indicates high expression. (b) Expression profiles of *CePIP1;1*, *-2;1*, and *-2;8* at different stages of leaf development. (c) Expression profiles of *CePIP1;1*, *-2;1*, and *-2;8* in mature leaves of diurnal fluctuation. Bars indicate SD (N = 3) and uppercase letters indicate difference significance tested following Duncan's one-way multiple-range post hoc ANOVA ($p< 0.01$). (Ce: *C. esculentus*; FPKM: Fragments per kilobase of exon per million fragments mapped; PIP: plasma membrane intrinsic protein)

shown in Fig. 3c. Generally, transcripts of all three genes in the day (8, 12, 16, and 20 h) were higher than that in the night (24 and 4 h). During the day, both *CePIP1;1* and *-2;8* exhibited an unimodal expression pattern that peaked at 12 h, whereas *CePIP2;1* possessed two peaks (8 and 16 h) and their difference was not significant. Nevertheless, transcripts of all three genes at 20 h (onset of night) were significantly lower than those at 8 h (onset of day) as well as 12 h. In the night, except for *CePIP2;1*, no significant difference was observed between the two stages for both *CePIP1;1* and *-2;8*. Moreover, their transcripts were comparable to those at 20 h (Fig. 3c).

Expression profiles of *CePIP* genes during tuber development

To reveal the expression patterns of *CePIP* genes during tuber development, three representative stages, i.e., 40 DAS (early swelling stage), 85 DAS (late swelling stage), and 120 DAS (mature stage), were first profiled using transcriptome data. As shown in Fig. 4a, except for rare expression of *CePIP1;2*, *-2;2*, *-2;9*, and *-2;10*, most genes exhibited a bell-like expression pattern peaking at 85 DAS, in contrast to a gradual decrease of *CePIP2;3* and *-2;8*. Notably, except for *CePIP2;4*, other genes were expressed considerably lower at 120 DAS than that at 40 DAS. For qRT-PCR confirmation of *CePIP1;1*, *-2;1*, and *-2;8*, seven stages were examined, i.e., 1, 5, 10, 15, 20, 25, and 35 DAI, which represent initiation, five stages of swelling, and maturation as described before[32]. As shown in Fig. 4b, two peaks were observed for all three genes, though their patterns were different. As for *CePIP1;1*, compared with the initiation stage (1 DAI), significant up-regulation was observed at the early swelling stage (5 DAI), followed by a gradual decrease except for the appearance of the second peak at 20 DAI, which is something different from transcriptome profiling. As for *CePIP2;1*, a sudden drop of transcripts first appeared at 5 DAI, then gradually increased until 20 DAI, which was followed by a gradual decrease at two late stages. The pattern of *CePIP2;8* is similar to *-1;1*, two peaks appeared at 5 and 20 DAI and the second peak was significantly lower than the first. The difference is that the second peak of *CePIP2;8* was significantly lower than the initiation stage. By contrast, the second peak (20 DAI) of *CePIP2;1* was significantly higher than that of the first one (1 DAI). Nevertheless, the expression patterns of both *CePIP2;1* and *-2;8* are highly consistent with transcriptome profiling.

Protein abundances of CePIP1;1, -2;1, and -2;8 in tubers are positively correlated with the moisture content during tuber development

Since protein abundance is not always in agreement with the transcript level, protein profiles of three dominant members (i.e. CePIP1;1, -2;1, and -2;8) during tuber development were further investigated. For this purpose, we first took advantage of available proteomic data to identify CePIP proteins, i.e., leaves, roots, and four stages of tubers (freshly harvested, dried, rehydrated for 48 h, and sprouted). As shown in Supplemental Fig. S2, all three proteins were identified in both leaves and roots, whereas CePIP1;1 and -2;8 were also identified in at least

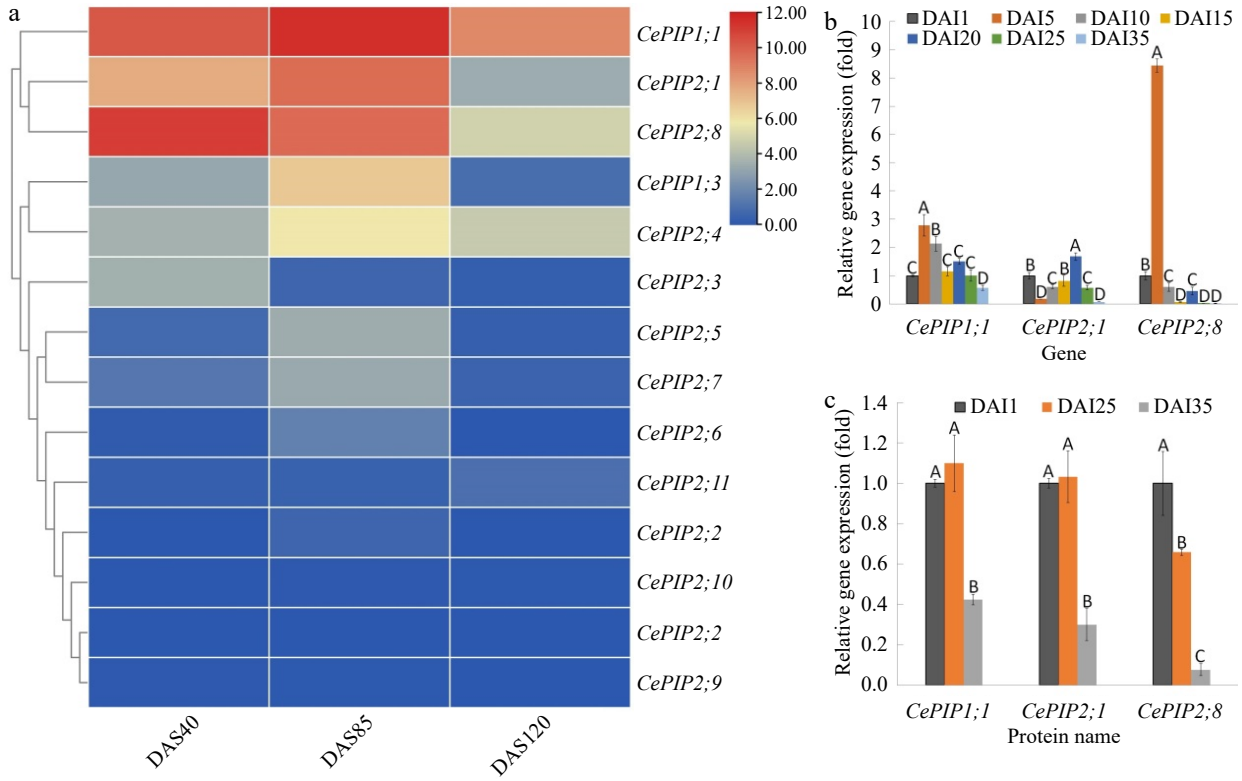


Fig. 4 Transcript and protein abundances of *CePIP* genes during tuber development. (a) Transcriptome-based expression profiling of 14 *CePIP* genes during tuber development. The heatmap was generated using the R package implemented with a row-based standardization. Color scale represents FPKM normalized $\log_2$ transformed counts, where blue indicates low expression and red indicates high expression. (b) qRT-PCR-based expression profiling of *CePIP1;1*, *-2;1*, and *-2;8* in seven representative stages of tuber development. (c) Relative protein abundance of CePIP1;1, -2;1, and -2;8 in three representative stages of tuber development. Bars indicate SD (N = 3) and uppercase letters indicate difference significance tested following Duncan's one-way multiple-range post hoc ANOVA ($p < 0.01$). (Ce: *C. esculentus*; DAI: days after tuber initiation; DAS: days after sowing; FPKM: Fragments per kilobase of exon per million fragments mapped; PIP: plasma membrane intrinsic protein).

one of four tested stages of tubers. Notably, all three proteins were considerably more abundant in roots, implying their key roles in root water balance.

To further uncover their profiles during tuber development, 4D-PRM-based protein quantification was conducted in three representative stages of tuber development, i.e., 1, 25, and 35 DAI. As expected, all three proteins were identified and quantified. In contrast to gradual decrease of CePIP2;8, both CePIP1;1 and -2;1 exhibited a bell-like pattern that peaked at 25 DAI, though no significant difference was observed between 1 and 25 DAI (Fig. 4c). The trends are largely in accordance with their transcription patterns, though the reverse trend was observed for CePIP2;1 at two early stages (Fig. 4b & Fig. 4c).

Subcellular localization analysis reveals that CePIP1;1, -2;1, and -2;8 function in the cell membrane

As predicted by WoLF PSORT, CePIP1;1, -2;1, and -2;8 may function in the cell membrane. To confirm the result, subcellular localization vectors named *p*NC-Cam1304-*CePIP1;1*, *p*NC-Cam1304-*CePIP2;1*, and *p*NC-Cam1304-*CePIP2;8* were further constructed. When transiently overexpressed in tobacco leaves, green fluorescence signals of all three constructs were confined to cell membranes, highly coinciding with red fluorescence signals of the plasma membrane marker HbPIP2;3-RFP (Fig. 5).

Discussion

Water balance is particularly important for cell metabolism and enlargement, plant growth and development, and stress responses[2,19]. As the name suggests, AQPs raised considerable interest for their high permeability to water, and plasma membrane-localized PIPs were proven to play key roles in transmembrane water transport between cells[1,18]. The first PIP was discovered in human erythrocytes, which was named CHIP28 or AQP1, and its homolog in plants was first characterized in Arabidopsis, which is known as RD28, PIP2c, or AtPIP2;3[3,7,53]. Thus far, genome-wide identification of *PIP* genes have been reported in a high number of plant species, including two model plants Arabidopsis and rice[10,11,13–17,54–56]. By contrast, little information is available on Cyperaceae, the third largest family within the monocot clade that possesses more than 5,600 species[57].

The tigernut genome encodes 14 *PIP* genes with three introns and gene expansion was mainly contributed by WGD and transposed/tandem duplications

Given the crucial roles of water balance for tuber development and crop production, in this study, tigernut, a representative Cyperaceae plant producing high amounts of oil in underground tubers[28,30,32], was employed to study *PIP* genes. A

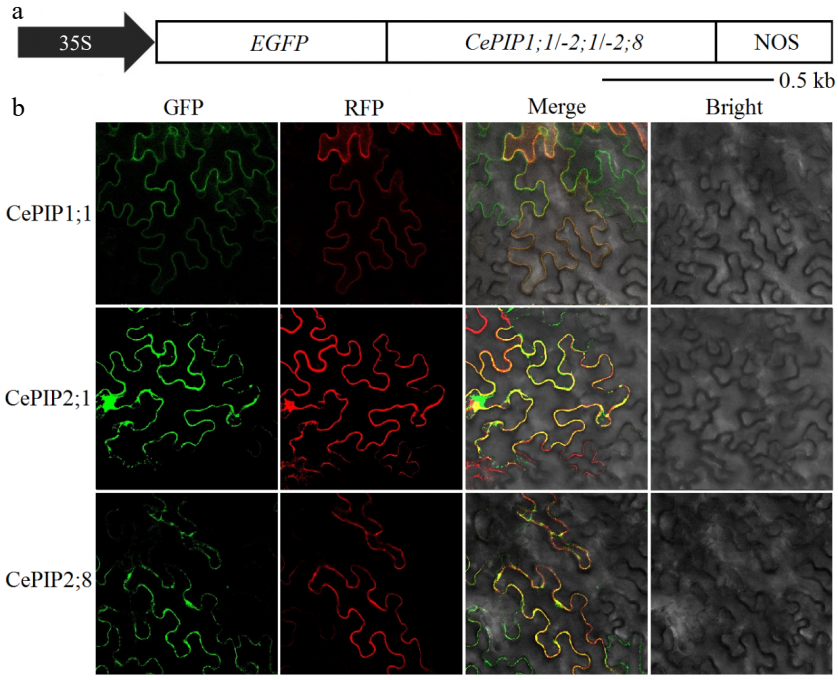


Fig. 5 (a) Schematic diagram of overexpressing constructs, (b) subcellular localization analysis of CePIP1;1, -2;1, and -2;8 in *N. benthamiana* leaves. (35S: cauliflower mosaic virus 35S RNA promoter; Ce: *C. esculentus*; EGFP: enhanced green fluorescent protein; kb: kilobase; NOS: terminator of the nopaline synthase gene; RFP: red fluorescent protein; PIP: plasma membrane intrinsic protein).

number of 14 *PIP* genes representing two phylogenetic groups (i.e., PIP1 and PIP2) or 12 orthogroups (i.e., PIP1A, PIP1B, PIP1C, PIP2A, PIP2B, PIP2C, PIP2D, PIP2E, PIP2F, PIP2G, PIP2H, and PIP2I) were identified from the tigernut genome. Though the family amounts are comparative or less than 13–21 present in Arabidopsis, cassava (*Manihot esculenta*), rubber tree (*Hevea brasiliensis*), poplar (*Populus trichocarpa*), *C. cristatella*, *R. breviuscula*, banana (*Musa acuminata*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), barley (*Hordeum vulgare*), and switchgrass (*Panicum virgatum*), they are relatively more than four to 12 found in eelgrass (*Zostera marina*), *Brachypodium distachyon*, foxtail millet (*Setaria italic*), *J. effuses*, *Aquilegia coerulea*, papaya (*Carica papaya*), castor been (*Ricinus communis*), and physic nut (*Jatropha curcas*) (Supplemental Table S4). Among them, *A. coerulea* represents a basal eudicot that didn't experience the γ WGD shared by all core eudicots[50], whereas eelgrass is an early diverged aquatic monocot that didn't experience the τ WGD shared by all core monocots[56]. Interestingly, though both species possess two PIP1s and two PIP2s, they were shown to exhibit complex orthologous relationships of 1:1, 2:2, 1:0, and 0:1 (Supplemental Table S5). Whereas AcPIP1;1/AcPIP1;2/ZmPIP1;1/ZmPIP1;2 and ZmPIP2;1/AcPIP2;1 belong to PIP1A and PIP2A identified in this study, AcPIP2;2 and ZmPIP2;2 belong to PIP2H and PIP2I, respectively (Supplemental Table S5), implying that the last common ancestor of monocots and eudicots possesses only one PIP1 and two PIP2s followed by clade-specific expansion. A good example is the generation of *AtPIP1;1* and *-2;6* from *AtPIP1;5* and *-2;1 via* the γ WGD, respectively[17].

In tigernut, extensive expansion of the PIP subfamily was contributed by WGD (2), transposed (2), tandem (5), and dispersed duplications (3). It's worth noting that, two transposed repeats (i.e., *CePIP1;1/-1;3* and *CePIP2;1/-2;8*) are shared by rice, implying their early origin that may be generated sometime after the split with the eudicot clade but before Cyperaceae-Poaceae divergence. By contrast, two WGD repeats (i.e., *CePIP1;1/-1;2* and *CePIP2;2/-2;4*) are shared by *C. cristatella*, *R. breviuscula*, and *J. effusus* but not rice and Arabidopsis, implying that they may be derived from WGDs that occurred sometime after Cyperaceae-Poaceae split but before Cyperaceae-Juncaceae divergence. The possible WGD is the one that was described in *C. littledalei*[58], though the exact time still needs to be studied. Interestingly, compared with Arabidopsis (1) and rice (2), tandem/proximal duplications played a more important role in the expansion of *PIP* genes in tigernut (5) as well as other Cyperaceae species tested (5–6), which were shown to be Cyperaceae-specific or even species-specific. These tandem repeats may play a role in the adaptive evolution of Cyperaceae species as described in a high number of plant species[14,41]. According to comparative genomics analyses, tandem duplicates experienced stronger selective pressure than genes formed by other modes (WGD, transposed duplication, and dispersed duplication) and evolved toward biased functional roles involved in plant self-defense[41].

As observed in most species such as Arabidopsis[10,14–17], *PIP* genes in all Cyperaceae and Juncaceae species examined in this study, i.e., tigernut, *C. cristatella*, *R. breviuscula*, and *J. effuses*, feature three introns with conserved positions. By contrast, zero to three introns was not only found in rice but also in other Poaceae species such as maize, sorghum, foxtail millet, switchgrass, *B. distachyon*, and barley[54,55], implying lineage/species-specific evolution.

CePIP1;1, *-2;1*, and *-2;8* are three dominant members that are constitutively expressed in most tissues

Despite the extensive expansion of *PIP* genes (*PIP2*) in tigernut even after the split with *R. breviuscula*, *CePIP1;1*, *-2;1*, and

-2;8 were shown to represent three dominant members in most tissues examined in this study, i.e., young leaf, mature leaf, sheath, rhizome, shoot apex, and tuber, though the situation in root is more complex. *CePIP1;1* was characterized as a transposed repeat of *CePIP1;3*, which represents the most expressed member in root. Moreover, its recent WGD repeat *CePIP1;2* was shown to be lowly expressed in most tested tissues, implying their divergence. The ortholog of *CePIP1;1* in rice is *OsPIP1;3* (*RWC-3*), which was shown to be preferentially expressed in roots, stems, and leaves, in contrast to constitutive expression of *OsPIP1;1* (*OsPIP1a*) and *-1;2*[59–61], two recent WGD repeats. Injecting the cRNA of *OsPIP1;3* into *Xenopus* oocytes could increase the osmotic water permeability by 2–3 times[60], though the activity is considerably lower than PIP2s such as OsPIP2;2 and -2;2[61–63]. Moreover, *OsPIP1;3* was shown to play a role in drought avoidance in upland rice and its overexpression in lowland rice could increase root osmotic hydraulic conductivity, leaf water potential, and relative cumulative transpiration at the end of 10 h PEG treatment[64]. *CePIP2;8* was characterized as a transposed repeat of *CePIP2;1*. Since their orthologs are present in both rice and Arabidopsis (Supplemental Table S3), the duplication event is more likely to occur sometime before monocot-eudicot split. Interestingly, their orthologs in rice, i.e., *OsPIP2;1* (*OsPIP2a*) and *-2;6*, respectively, are also constitutively expressed[61], implying a conserved evolution with similar functions. When heterologously expressed in yeast, *OsPIP2;1* was shown to exhibit high water transport activity[62,64–66]. Moreover, root hydraulic conductivity was decreased by approximately four folds in *OsPIP2;1* RNAi knockdown rice plants[64]. The water transport activity of OsPIP2;6 has not been tested, however, it was proven to be an H_2O_2 transporter that is involved in resistance to rice blast[61]. More work especially transgenic tests may improve our knowledge of the function of these key *CePIP* genes.

Transcription of *CePIP1;1*, *-2;1*, and *-2;8* in leaves is subjected to development and diurnal regulation

Leaf is a photosynthetic organ that regulates water loss through transpiration. In tigernut, PIP transcripts in leaves were mainly contributed by *CePIP1;1*, *-2;1*, and *-2;8*, implying their key roles. During leaf development, in contrast to gradual decrease of *CePIP1;1* and *-2;8* transcripts in three stages (i.e. young, mature, and senescing) examined in this study, *CePIP2;1* peaked in mature leaves. Their high abundance in young leaves is by cell elongation and enlargement at this stage, whereas upregulation of *CePIP2;1* in mature leaves may inform its possible role in photosynthesis[67]. Thus far, a high number of CO_2 permeable PIPs have been identified, e.g., AtPIP2;1, HvPIP2;1, HvPIP2;2, HvPIP2;3, HvPIP2;5, and SiPIP2;7[68–70]. Moreover, in mature leaves, *CePIP1;1*, *-2;1*, and *-2;8* were shown to exhibit an apparent diurnal fluctuation expression pattern that was expressed more in the day and usually peaked at noon, which reflects transpiration and the fact that *PIP* genes are usually induced by light[11,71–73]. In rice, *OsPIP2;4* and *-2;5* also showed a clear diurnal fluctuation in roots that peaked at 3 h after the onset of light and dropped to a minimum 3 h after the onset of darkness[11]. Notably, further studies showed that temporal and dramatic induction of *OsPIP2;5* around 2 h after light initiation was triggered by transpirational demand but not circadian rhythm[74].

mRNA and protein abundances of *CePIP1;1*, *-2;1*, and *-2;8* are positively correlated with the moisture content during tuber development

As an oil-bearing tuber crop, the main economic goal of tigernut cultivation is to harvest underground tubers, whose development is highly dependent on water available[32,75]. According to previous studies, the moisture content of immature tigernut tubers maintains more than 80.0%, followed by a seed-like dehydration process with a drop of water content to less than 50% during maturation[28,32]. Thereby, the water balance in developmental tubers must be tightly regulated. Like leaves, the majority of PIP transcripts in tubers were shown to be contributed by *CePIP1;1*, *-2;1*, and *-2;8*, which was further confirmed at the protein level. In accordance with the trend of water content during tuber development, mRNA, and protein abundances of *CePIP1;1*, *-2;1*, and *-2;8* in initiation and swelling tubers were considerably higher than that at the mature stage. High abundances of *CePIP1;1*, *-2;1*, and *-2;8* at the initiation stage reflects rapid cell division and elongation, whereas upregulation of *CePIP1;1* and *-2;1* at the swelling stage is in accordance with cell enlargement and active physiological metabolism such as rapid oil accumulation[28,30]. At the mature stage, downregulation of PIP transcripts and protein abundances resulted in a significant drop in the moisture content, which is accompanied by the significant accumulation of late embryogenesis-abundant proteins[23,32]. The situation is highly distinct from other tuber plants such as potato (*Solanum tuberosum*), which may contribute to the difference in desiccation resistance between two species[32,76]. It's worth noting that, in one study, CePIP2;1 was not detected in any of the four tested stages, i.e., freshly harvested, dried, rehydrated for 48 h, and sprouted tubers[23]. By contrast, it was quantified in all three stages of tuber development examined in this study, i.e., 1, 25, and 35 DAI (corresponding to freshly harvested tubers), which represent initiation, swelling, and maturation. One possible reason is that the protein abundance of CePIP2;1 in mature tubers is not high enough to be quantified by nanoLC-MS/MS, which is relatively less sensitive than 4D-PRM used in this study[30,46]. In fact, nanoLC-MS/MS-based proteomic analysis of 30 samples representing six tissues/stages only resulted in 2,257 distinct protein groups[23].

Taken together, our results imply a key role of CePIP1;1, -2;1, and -2;8 in tuber water balance, however, the mechanism underlying needs to be further studied, e.g., posttranslational modifications, protein interaction patterns, and transcriptional regulators.

Conclusions

To our knowledge, this is the first genome-wide characterization of *PIP* genes in tigernut, a representative Cyperaceae plant with oil-bearing tubers. Fourteen *CePIP* genes representing two phylogenetic groups or 12 orthogroups are relatively more than that present in two model plants rice and Arabidopsis, and gene expansion was mainly contributed by WGD and transposed/tandem duplications, some of which are lineage or even species-specific. Among these genes, *CePIP1;1*, *-2;1*, and *-2;8* have evolved to be three dominant members that are constitutively expressed in most tissues, including leaf and tuber. Transcription of these three dominant members in leaves are subjected to development and diurnal regulation, whereas in

tubers, their mRNA and protein abundances are positively correlated with the moisture content during tuber development. Moreover, their plasma membrane-localization was confirmed by subcellular localization analysis, implying that they may function in the cell membrane. These findings shall not only provide valuable information for further uncovering the mechanism of tuber water balance but also lay a solid foundation for genetic improvement by regulating these key PIP members in tigernut.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design, supervision: Zou Z; analysis and interpretation of results: Zou Z, Zheng Y, Xiao Y, Liu H, Huang J, Zhao Y; draft manuscript preparation: Zou Z, Zhao Y. All authors reviewed the results and approved the final version of the manuscript.

Data availability

All the relevant data is available within the published article.

Acknowledgments

This work was supported by the Hainan Province Science and Technology Special Fund (ZDYF2024XDNY171 and ZDYF2024XDNY156), China; the National Natural Science Foundation of China (32460342, 31971688 and 31700580), China; the Project of Sanya Yazhou Bay Science and Technology City (SCKJ-JYRC-2022-66), China. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/tp-0024-0030)

Dates

Received 1 May 2024; Revised 26 June 2024; Accepted 15 July 2024; Published online 22 August 2024

References

1. Agre P. 2006. The aquaporin water channels. *Proceedings of the American Thoracic Society* 3:5–13
2. Tyerman SD, McGaughey SA, Qiu J, Yool AJ, Byrt CS. 2021. Adaptable and multifunctional ion-conducting aquaporins. *Annual Review of Plant Biology* 72:703–36
3. Preston GM, Carroll TP, Guggino WB, Agre P. 1992. Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* 256:385–87
4. Ishibashi K, Tanaka Y, Morishita Y. 2023. Evolutionary overview of aquaporin superfamily. In *Aquaporins. Advances in Experimental Medicine and Biology*, ed. Yang B. vol 1398. Singapore: Springer. pp. 81–98. DOI: 10.1007/978-981-19-7415-1_6
5. Törnroth-Horsefield S, Wang Y, Hedfalk K, Johanson U, Karlsson M, et al. 2006. Structural mechanism of plant aquaporin gating. *Nature* 439:688–94
6. Fu D, Libson A, Miercke LJ, Weitzman C, Nollert P, et al. 2000. Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* 290:481–86
7. Sui H, Han BG, Lee JK, Walian P, Jap BK. 2001. Structural basis of water specific transport through the AQP1 water channel. *Nature* 414:872–78
8. Byrt CS, Zhao M, Kourghi M, Bose J, Henderson SW, et al. 2017. Non-selective cation channel activity of aquaporin AtPIP2;1 regulated by Ca^{2+} and pH. *Plant Cell and Environment* 40:802–15
9. Wang H, Schoebel S, Schmitz F, Dong H, Hedfalk K. 2020. Characterization of aquaporin-driven hydrogen peroxide transport. *Biochimica et Biophysica Acta (BBA) - Biomembranes,* 1862:183065.
10. Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjövall S, et al. 2001. The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants. *Plant Physiology* 126:1358–69
11. Sakurai J, Ishikawa F, Yamaguchi T, Uemura M, Maeshima M. 2005. Identification of 33 rice aquaporin genes and analysis of their expression and function. *Plant and Cell Physiology* 46:1568–77
12. Abascal F, Irisarri I, Zardoya R. 2014. Diversity and evolution of membrane intrinsic proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1840:1468–81
13. Zou Z, Gong J, An F, Xie G, Wang J, et al. 2015. Genome-wide identification of rubber tree (*Hevea brasiliensis* Muell. Arg.) aquaporin genes and their response to ethephon stimulation in the laticifer, a rubber-producing tissue. *BMC Genomics* 16:1001
14. Zou Z, Gong J, Huang Q, Mo Y, Yang L, et al. 2015. Gene structures, evolution, classification and expression profiles of the aquaporin gene family in castor bean (*Ricinus communis* L.). *PLoS One* 10:e0141022
15. Zou Z, Yang L, Gong J, Mo Y, Wang J, et al. 2016. Genome-wide identification of *Jatropha curcas* aquaporin genes and the comparative analysis provides insights into the gene family expansion and evolution in *Hevea brasiliensis*. *Frontiers in Plant Science* 7:395
16. Zou Z, Yang J. 2019. Genome-wide comparison reveals divergence of cassava and rubber aquaporin family genes after the recent whole-genome duplication. *BMC Genomics* 20:380
17. Zou Z, Zheng Y, Xie Z. 2023. Analysis of *Carica papaya* informs lineage-specific evolution of the aquaporin (AQP) family in Brassicales. *Plants* 12:3847
18. Wudick MM, Luu DT, Maurel C. 2009. A look inside: localization patterns and functions of intracellular plant aquaporins. *New Phytologist* 184:289–302
19. Chaumont F, Tyerman SD. 2014. Aquaporins: highly regulated channels controlling plant water relations. *Plant Physiology* 164:1600–18
20. Qiao X, Zheng Y, Yang J, Zeng C, Zou Z. 2022. Gene cloning, subcellular localization and multimerization analysis of *HbPIP1;1* from *Hevea brasiliensis*. *Chinese Journal of Tropical Crops* 43:2405–12
21. Zheng YJ, Chang LL, Zhao YG, Zeng CY, Zou Z. 2024. Molecular cloning and characterization of *CePIP1;1*, an aquaporin gene highly abundant in *Cyperus esculentus* tubers. *Chinese Journal of Tropical Crops* 45:894–901
22. Zou Z, Zheng Y, Qiao X, Yang J. 2024. Subcellular localization and multimerization analyses of HbPIP2;3, an efficient water transporter from *Hevea brasiliensis*. *Chinese Journal of Tropical Crops,* 45:443–49
23. Niemeyer PW, Irisarri I, Scholz P, Schmitt K, Valerius O, et al. 2022. A seed-like proteome in oil-rich tubers. *The Plant Journal* 112:518–34
24. Xiao Y, Zou Z, Zhao Y, Guo A, Zhang L. 2022. Molecular cloning and characterization of an acetolactate synthase gene (*CeALS*) from tigernut (*Cyperus esculentus* L.). *Biotechnology Bulletin* 38:184–92
25. Zou Z, Xiao YH, Zhang L, Zhao YG. 2023. Analysis of *Lhc* family genes reveals development regulation and diurnal fluctuation expression patterns in *Cyperus esculentus*, a Cyperaceae plant. *Planta* 257:59

26. Zou Z, Xiao Y, Zhang L, Zhao Y. 2023. Cloning and characterization of *CeEPSPS*, a gene encoding 5-enolpyruvylshikimate-3-phosphate synthase from tigernut (*Cyperus esculentus* L.). *Chinese Journal of Tropical Crops* 44(1):26−34
27. Zou Z, Zheng Y, Chang L, Zou L, Zhang L, et al. 2024. TIP aquaporins in *Cyperus esculentus*: genome-wide identification, expression profiles, subcellular localizations, and interaction patterns. *BMC Plant Biology* 24:298
28. Turesson H, Marttila S, Gustavsson KE, Hofvander P, Olsson ME, et al. 2010. Characterization of oil and starch accumulation in tubers of *Cyperus esculentus* var. sativus (Cyperaceae): A novel model system to study oil reserves in nonseed tissues. *American Journal of Botany* 97:1884−93
29. Xu S, Zou Z, Xiao Y, Zhang L, Kong H, et al. 2022. Cloning and functional characterization of *CeWRI1*, a gene involved in oil accumulation from tigernut (*Cyperus esculentus* L.) tubers. *Chinese Journal of Tropical Crops* 43:923−29
30. Zou Z, Zheng Y, Zhang Z, Xiao Y, Xie Z, et al. 2023. Molecular characterization of *oleosin* genes in *Cyperus esculentus*, a Cyperaceae plant producing oil in underground tubers. *Plant Cell Reports* 42:1791−808
31. Zou Z, Zhao Y, Zhang L, Kong H, Guo Y, et al. 2021. Single-molecule real-time (SMRT)-based full-length transcriptome analysis of tigernut (*Cyperus esculentus* L.). *Chinese Journal of Oil Crop Sciences* 43:229−35
32. Zou Z, Zhao Y, Zhang L, Xiao Y, Guo A. 2022. Analysis of *Cyperus esculentus SMP* family genes reveals lineage-specific evolution and seed desiccation-like transcript accumulation during tuber maturation. *Industrial Crops and Products* 187:115382
33. Bai X, Chen T, Wu Y, Tang M, Xu ZF. 2021. Selection and validation of reference genes for qRT-PCR analysis in the oil-rich tuber crop tiger nut (*Cyperus esculentus*) based on transcriptome data. *International Journal of Molecular Sciences* 22:2569
34. Zhao X, Yi L, Ren Y, Li J, Ren W, et al. 2023. Chromosome-scale genome assembly of the yellow nutsedge (*Cyperus esculentus*). *Genome Biology and Evolution* 15:evad027
35. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:3389−402
36. Planta J, Liang YY, Xin H, Chansler MT, Prather LA, et al. 2022. Chromosome-scale genome assemblies and annotations for Poales species *Carex cristatella, Carex scoparia, Juncus effusus*, and *Juncus inflexus*. *G3 Genes|Genomes|Genetics* 12:jkac211
37. Hofstatter PG, Thangavel G, Lux T, Neumann P, Vondrak T, et al. 2022. Repeat-based holocentromeres influence genome architecture and karyotype evolution. *Cell* 185:3153−68
38. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30:2725−29
39. Chen C, Wu Y, Li J, Wang X, Zeng Z, et al. 2023. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Molecular Plant* 16:1733−42
40. Zou Z, Yang J, Zhang X. 2019. Insights into genes encoding respiratory burst oxidase homologs (RBOHs) in rubber tree (*Hevea brasiliensis* Muell. Arg.). *Industrial Crops and Products* 128:126−39
41. Qiao X, Li Q, Yin H, Qi K, Li L, et al. 2019. Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. *Genome Biology* 20:38
42. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24:1586−91
43. Persson E, Sonnhammer ELL. 2022. InParanoid-DIAMOND: faster orthology analysis with the InParanoid algorithm. *Bioinformatics* 38:2918−19
44. Zou Z, Xie G, Yang L. 2017. Papain-like cysteine protease encoding genes in rubber (*Hevea brasiliensis*): Comparative genomics, phylogenetic and transcriptional profiling analysis. *Planta* 246:999−1018
45. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature Methods* 5:621−28
46. Wang C, Chen L, Cai ZC, Chen C, Liu Z, et al. 2020. Comparative proteomic analysis reveals the molecular mechanisms underlying the accumulation difference of bioactive constituents in *Glycyrrhiza uralensis* fisch under salt stress. *Journal of Agricultural and Food Chemistry* 68:1480−93
47. Maurel C, Verdoucq L, Luu DT, Santoni V. 2008. Plant aquaporins: membrane channels with multiple integrated functions. *Annual Review of Plant Biology* 59:595−624
48. Tournaire-Roux C, Sutka M, Javot H, Gout E, Gerbeau P, et al. 2003. Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins. *Nature* 425:393−97
49. Chaumont F, Barrieu F, Jung R, Chrispeels MJ. 2000. Plasma membrane intrinsic proteins from maize cluster in two sequence subgroups with differential aquaporin activity. *Plant Physiology* 122:1025−34
50. Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers JE, McKain MR, et al. 2012. A genome triplication associated with early diversification of the core eudicots. *Genome Biology* 13:R3
51. Jiao Y, Li J, Tang H, Paterson AH. 2014. Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. *The Plant Cell* 26:2792−802
52. Bowers JE, Chapman BA, Rong J, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422:433−8
53. Daniels MJ, Mirkov TE, Chrispeels MJ. 1994. The plasma membrane of *Arabidopsis thaliana* contains a mercury-insensitive aquaporin that is a homolog of the tonoplast water channel protein TIP. *Plant Physiology* 106:1325−33
54. Chaumont F, Barrieu F, Wojcik E, Chrispeels MJ, Jung R. 2001. Aquaporins constitute a large and highly divergent protein family in maize. *Plant Physiology* 125:1206−15
55. Azad AK, Ahmed J, Alum MA, Hasan MM, Ishikawa T, et al. 2016. Genome-wide characterization of major intrinsic proteins in four grass plants and their non-aqua transport selectivity profiles with comparative perspective. *PLoS One* 11:e0157735
56. Shivaraj SM, Deshmukh R, Bhat JA, Sonah H, Bélanger RR. 2017. Understanding aquaporin transport system in eelgrass (*Zostera marina* L.), an aquatic plant species. *Frontiers in Plant Science* 8:1334
57. Govaerts R, Simpson DA, Goetghebeur P, Wilson KL, Egorova T, et al. 2007. World checklist of Cyperaceae. The Board of Trustees of the Royal Botanic Gardens, Kew
58. Can M, Wei W, Zi H, Bai M, Liu Y, et al. 2020. Genome sequence of *Kobresia littledalei*, the first chromosome-level genome in the family Cyperaceae. *Scientific Data* 7:175
59. Malz S, Sauter M. 1999. Expression of two PIP genes in rapidly growing internodes of rice is not primarily controlled by meristem activity or cell expansion. *Plant Molecular Biology* 40:985−95
60. Li L, Li S, Tao Y, Kitagawa Y. 2000. Molecular cloning of a novel water channel from rice: its products expression in *Xenopus* oocytes and involvement in chilling tolerance. *Plant Science* 154:43−51
61. Li G, Han J, Yi C, Luo H, Wang Y, et al. 2024. Global characterization of OsPIP aquaporins reveals that the H_2O_2 transporter OsPIP2;6 increases resistance to rice blast. *The Crop Journal* 12:102−9
62. Ding L, Uehlein N, Kaldenhoff R, Guo S, Zhu Y, et al. 2019. Aquaporin *PIP2;1* affects water transport and root growth in rice (*Oryza sativa* L.). *Plant Physiology and Biochemistry* 139:152−60
63. Liu S, Fukumoto T, Gena P, Feng P, Sun Q, et al. 2020. Ectopic expression of a rice plasma membrane intrinsic protein (OsPIP1;3) promotes plant growth and water uptake. *The Plant Journal* 102:779−96
64. Bai J, Wang X, Yao X, Chen X, Lu K, et al. 2021. Rice aquaporin OsPIP2;2 is a water-transporting facilitator in relevance to drought-tolerant responses. *Plant Direct* 5:e338

65. Lian HL, Yu X, Ye Q, Ding XS, Kitagawa Y, et al. 2004. The role of aquaporin RWC3 in drought avoidance in rice. *Plant and Cell Physiology* 45:481−89
66. Sakurai J, Ahamed A, Murai M, Maeshima M, Uemura M. 2008. Tissue and cell-specific localization of rice aquaporins and their water transport activities. *Plant and Cell Physiology* 49:30−39
67. Heinen RB, Ye Q, Chaumont F. 2009. Role of aquaporins in leaf physiology. *Journal of Experimental Botany* 60:2971−85
68. Mori IC, Rhee J, Shibasaka M, Sasano S, Kaneko T, et al. 2014. CO_2 transport by PIP2 aquaporins of barley. *Plant and Cell Physiology* 55:251−7
69. Wang C, Hu H, Qin X, Zeise B, Xu D, et al. 2016. Reconstitution of CO_2 regulation of SLAC1 anion channel and function of CO_2-permeable PIP2;1 aquaporin as CARBONIC ANHYDRASE4 interactor. *The Plant Cell* 28:568−82
70. Ermakova M, Osborn H, Groszmann M, Bala S, Bowerman A, et al. 2021. Expression of a CO_2-permeable aquaporin enhances mesophyll conductance in the C_4 species *Setaria viridis*. *eLife* 10:e70095
71. Baaziz KB, Lopez D, Rabot A, Combes D, Gousset A, et al. 2012. Light-mediated K_{leaf} induction and contribution of both the PIP1s and PIP2s aquaporins in five tree species: Walnut (*Juglans regia*) case study. *Tree Physiology* 32:423−34
72. Perez-Martin A, Michelazzo C, Torres-Ruiz JM, Flexas J, Fernández JE, et al. 2014. Regulation of photosynthesis and stomatal and mesophyll conductance under water stress and recovery in olive trees: Correlation with gene expression of carbonic anhydrase and aquaporins. *Journal of Experimental Botany* 65:3143−56
73. An F, Zou Z, Cai X, Wang J, Rookes J, et al. 2015. Regulation of *HbPIP2;3*, a latex-abundant water transporter, is associated with latex dilution and latex yield in rubber tree (*Hevea brasiliensis* MuellArg.). *PLoS One* 10:e0125595
74. Sakurai-Ishikawa J, Murai-Hatano M, Hayashi H, Ahamed A, Fukushi K, et al. 2011. Transpiration from shoots triggers diurnal changes in root aquaporin expression. *Plant, Cell & Environment* 34:1150−63
75. Yang X, Niu L, Zhang Y, Ren W, Yang C, et al. 2022. Morpho-agronomic and biochemical characterization of accessions of tiger nut (*Cyperus esculentus*) grown in the north temperate zone of China. *Plants* 11:923
76. Delahaie J, Hundertmark M, Bove J, Leprince O, Rogniaux H, et al. 2013. LEA polypeptide profiling of recalcitrant and orthodox legume seeds reveals ABI3-regulated LEA protein abundance linked to desiccation tolerance. *Journal of Experimental Botany* 64:4559−73